Contingent Valuation of Competing Public Sector Programmes: An Experiment of Single versus Joint Evaluation

Cam Donaldson, Stephane Luchini, Jean-Paul Moatti, Christel Protière

GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

Document de Travail
n°2008-27

CONTINGENT VALUATION OF COMPETING PUBLIC SECTOR PROGRAMMES: AN EXPERIMENT OF SINGLE VERSUS JOINT EVALUATION

Cam DONALDSON
Stéphane LUCHINI
Christel PROTIERE
Jean-Paul MOATTI

April 2008

Contingent Valuation of Competing Public Sector Programmes: An Experiment of Single versus Joint Evaluation

Cam Donaldson
Institute of Health and Society
University of Newcastle

Stéphane Luchini
CNRS-GREQAM UMR 6579-IDEP

Jean-Paul Moatti
UMR 912 INSERM - IRD
Université de la Méditerranée

Christel Protière
UMR 912 INSERM - IRD
Université de la Méditerranée

May 2008

Abstract

In this paper, we compare single and joint evaluation (JE) of competing public sector programmes in a contingent valuation exercise. Using survey data aimed at evaluating WTP for cancer interventions ($n = 2628$), we disantangle two types of effects of JE: informational effects and sequence effects. By the former, we mean: by presenting different programmes to respondents, they will acquire more information on each programme than they would if each programme was valued in isolation. Sequence effects are underisable and induced by the JE exercise itself: changing the order of the valuation sequence induces different WTP values. Our results show that there are informational effects but no sequence effects. We therefore argue that JE approaches can be added to the armoury of techniques aimed at designing better survey instruments in a way that induces informational effects without incurring problems of sequencing.

Keywords: Single Evaluation, Joint Evaluation, Willingness to Pay
Contingent Valuation, Priorities Setting

JEL Classification: H4, I1

Corresponding author: Stephane Luchini
GREQAM
2 rue de la Charité
F-13002 Marseilles, France.

Phone: +33 (0) 491 140 789
e-mail: stephane.luchini@univmed.fr

1 Introduction

In many areas of public sector activity, several programmes will be competing for available funds at the margin. With no price data revealed from real choices to guide society to an efficient equilibrium, contingent valuation has emerged as the main method for completing the 'market' by aiding decisions on allocations across such programmes. In general, this process of completion is achieved through appropriately targeted groups of respondents (whether it be particular user groups or members of the general public) being asked about their maximum willingness to pay (WTP) for each of the programmes at stake.

This raises the question, however, of whether the process of completion is better achieved through a separate evaluation (SE) process, whereby different groups of respondents are each asked to value just one of the rival programmes, or a joint evaluation (JE) process, whereby all respondents value each of the rivals.

Although there may be obvious practical reasons of cost in favour of JE, it poses both theoretical and econometric challenges. On the former, standard economic theory might state that WTP values arising from both SE and JE procedures would be equivalent. However, this would ignore theoretical developments related to reference points and other explanations of preference reversals. Despite appearing as challenges to JE, it may be that such aspects are in fact better exploited in the JE situation, especially in the presence of imperfectly informed consumers (or respondents). On the latter, more thorough econometric procedures need to be used to test whether JE is indeed different to, and an improvement on, SE.

This paper reports the results of an experiment aimed at addressing these two issues. In the following section, a brief review of the literature and theoretical exposition of the potential differences between the SE and JE approaches reveals more clearly the need for the experiment conducted. This is followed by a description of the survey, experimental design and econometric testing procedures used to compare SE and JE. The results, discussion and conclusions then follow respectively.

2 Background and theory

Already reviewed by Bazerman et al. (1999), much previous literature comparing SE and JE has, for the most part, sought to determine whether preference reversals arise in situations in which one choice conveys a financial gain to the respondent, but in a scenario where s/he is worse off than his/her peers, whilst the alternative conveys a smaller financial gain, but with peer parity. Of course, other studies have also used JE and SE comparisons to examine similar sets of choices between consumer goods and part-whole biases (Hsee 1996; Hsee 1998; Hoehn and Randall 1989). Sequencing effects, too, have been thoroughly examined using JE approaches (Carson et al. 1998).

However, in the study reported here, SE and JE approaches are used to compare health goods for which there would be no prior expectation on the part of the analyst on preference orderings, whether viewed by the respondent from the perspective of efficiency or equality. This is similar to work examining whether people would express their support for government intervention in the case of very diverse and clearly mutually exclusive 'goods' such as treating skin cancer in farm workers or preserving a rare species of animal (Kahneman and Ritov 1994).

In the context of public sector decision making, where several alternatives are competing for limited funds, it would seem natural to expose a sample of the relevant constituency of respondents to descriptions of all of the alternatives requiring assessment and to ask all respondents to value each alternative. In this respect, JE would seem to offer several practical and theoretical advantages. First, it places the respondent in the same position as the decision maker in having to choose directly between several alternatives on offer at any one point in time. Secondly, it mimics a market situation in which the consumer is continuously weighing up the benefits of different goods.

This latter point indicates a further strength; that is that JE also avoids the more complex cognitive task associated with SE in which the respondent has to generate internal referents, all of which may also differ across such respondents, when assessing the single option put to them for valuation (Kahneman and Ritov 1994). It could be argued that providing respondents with a clearer idea of the alternatives to be compared in an evaluation, through a JE approach, will standardise the reference point. By focussing on what Bazerman et al. (1999) describe as "the comparison set for evaluation", a more discriminatory result is likely to arise than simply asking different groups to value just one programme each. For example, in the case of the study reported on in this paper, in which different cancer programmes are assessed, respondents valuing one programme may simply compare this to no programme at all, or even worse, either register a token amount as a 'vote' for cancer or use an unobserved reference point. The 'noise' generated by such behaviour has led in the past to results of several WTP studies not being able to discriminate between options evaluated and has led to a call for JE approaches (Johannesson and Fegeberg 1992; Ryan et al. 1997; Donaldson et al. 1995).

More formally, according to standard consumer theory, one would expect that values elicited under each procedure would be equal. Consider a simple model where an individual derives utility from his/er income y and his/er health h and consider further that his/er health depends upon health care provision. When a particular health programme, say A, is provided in a SE exercise, his/er WTP for this health programme is usually defined as

$$U(y, h_0) = U(y - WTP_A^s, h(A|_A)) \tag{1}$$

where h_0 is individual's health state with the current level of care, the subscript s indicates that the WTP is elicited in a SE exercise and $h(A|A)$ is the health state reached when A is provided. With this formulation we also want to indicate that the agent's evaluation of programme A on her health will reflect only the consideration of this particular health programme as in standard SE exercises.

Now consider that two mutually exclusive programmes, say A and B[1], are provided in a JE exercise and let us focus on WTP for programme A. The respondent is asked to value programme A, but before that, s/he has also been provided with the description of the other programme B. As in the previous scenario, one can derive his WTP for programme A so that

$$U(y, h_0) = U(y - WTP_A^d, h(A|_{A,B})) \tag{2}$$

where the subscript d indicates that agent's WTP for programme A is obtained in a JE of two mutually exclusive health programmes. For the sake of simplicity, consider that his/er utility is additively separable, so that $U(y, h) = g(y) + u(h)$, where $g(.)$ and $u(.)$ are standard monotonic increasing concave functions. This leads to

$$\begin{aligned} g(y) - g(y - WTP_A^s) &= u(h(A|_A)) - u(h_0) \quad \text{and} \\ g(y) - g(y - WTP_A^d) &= u(h(A|_{A,B})) - u(h_0) \end{aligned} \tag{3}$$

The difference between the two equations is

$$g(y - WTP_A^d) - g(y - WTP_A^s) = u(h(A|_A)) - u(h(A|_{A,B})). \tag{4}$$

[1]By mutually exclusive programmes, we consider health programmes that cannot *technically* be provided jointly such as, for instance, two surgical techniques, or health programmes that are made *experimentally* mutually exclusive – that is, one or the other programme is said to be available but not both. In doing so, we thus avoid potential substitution or complementary effects that would put on an additional layer of difficulties for the experimental study. We of course do not deny their relevance for decision-making in health care.

It is obvious that the WTP for programme A is identical in the SE and JE exercise if $u(h(A|_A)) = u(h(A|_{A,B}))$.[2] In other words, this would mean that the addition of another programme B to the evaluation exercise does not have any effect on the agent's evaluation of A. This is indeed the case of a standard perfectly informed and rational individual. This will be the null hypothesis that we are going to test.

If it turns out that this hypothesis does not hold and people will respond differently in SE and JE evaluations, then one can argue that, first, the presence or not of another programme B has an "informational effect" on people's valuations. By an informational effect we mean: by presenting programme A to individuals, they will come to learn what effect this programme has on their health. The same is true when they are presented with programme B. However, in addition to this "direct" information, individuals will acquire two further "indirect" types of "information"; namely, first, what is the difference on health of programme A and B and, second, by understanding this difference, they will also understand better the respective impact of each of the programmes on their health (Protière et al. 2004). If it is assumed that health care consumers are not well informed, it would follow that they are more likely, in the SE situation, to choose different and even inappropriate reference points, relative to the JE situation in which a relevant alternative is presented and for which a valuation is also asked for. Second, the null hypothesis may not hold because of differences between SE and JE induced by the JE exercise itself. This would amount to "sequence effects" as described by Payne et al. (2000) and Stewart et al. (2002). That is, changing the order in which the WTP is elicited for different programmes will have an impact on people's valuations.

Econometrically, there are techniques that should be used to estimate WTP values obtained in the context of a JE procedure and, thus, test more thoroughly the above hypotheses. In most previous studies, values have been estimated as though they are independent of each other (Donaldson et al. 1997; Olsen and Donaldson 1998). However, such estimations do not account for the JE procedure exogenously providing a reference structure for respondents which may influence their WTP valuations of each programme. This issue was investigated in an earlier study in which WTP values, several of which were derived from each respondent, were estimated as though independent (using ordinary least squares) and within a simultaneous equation framework (using seemingly unrelated regressions, or SUR) (Luchini et al. 2003). The main results of

[2]One could also think that the respondent's valuation of health given the current level of care, $u(h_0)$, depends either on the presence of A only or on A and B together in the evaluation exercise. That would lead to comparison of variations of utility instead of levels but would not change the conclusion.

this study were that independent estimation led to misspecification and that the use of SUR improved the efficiency of WTP estimates. However, the main conclusion was that this should be tested further in an experimental situation in which respondents would be randomly allocated to a SE or JE elicitation procedure. It is the result of such an experiment that is reported in this paper and which can further inform us as to these two important issues. In the following section of the paper, this experiment is described along with a more formal outline of the econometric analysis strategy.

3 Survey design and data

In a survey dealing with "knowledge and behaviors toward cancer", conducted during April-June 2005 by the National Institute of Prevention and Education for Health (INPES), 4046 members of the general population, 16 years of age and over, were interviewed by phone (Guilbert et al. 2006). One part of the questionnaire was dedicated to the contingent valuation study reported on here.

In this study WTP values were elicited for three health programmes on colorectal cancer, this being the second most diagnosed cancer in western nations (Ferlay et al. 2004). The three health programmes to be valued each represent different aspects of care within cancer; a "screening programme", a "new treatment programme" and a "rehabilitation programme" which we later refer to as programmes A, B and C respectively (see Appendix A for a complete description of each).

In order to compare SE and JE, some randomly-selected respondents had to perform a WTP evaluation for only one of these three programmes while others had to perform a JE of two of these three programmes. A comparison of WTP values obtained in SE and JE will allow us to test our null hypothesis of no difference between SE and JE. Moreover, if it turns out that some differences appear, all combinations of two programmes were evaluated in a different sequence in order to distinguish between the informational effect assumption and the sequence effect assumption. There were therefore three SE exercises of programmes A, B and C and six JE exercises: AB, BA, BC, CB, AC and CA.

More specifically, after explaining the purpose of the exercise, respondents were presented with description(s) of the programme(s) to be valued. Note that our design differs from Payne et al. (2000). In their study, Payne et al. (2000) informed respondents that they would evaluate five different programmes. However, information on each programme was not provided prior to the evaluation exercise but sequentially. Namely, the information on a particular programme was provided when the respondent was asked his WTP for the programme. Such a methodological choice can undermine the distinction between informational effects and sequence effects without any additional assumptions.

In the JE exercise, respondents were first asked to rank the two health programmes submitted for valuation. Then, WTP values were elicited using a two-step procedure. Firstly, the respondent was asked whether s/he would be willing to pay, and, if so, what would be the maximum amount in terms of a yearly additional contribution to the Sickness Fund of the French Social Security, the main financing scheme for health care in France. If respondents were not willing to pay, they were asked why using a pre-coded list of six alternative reasons. This allowed us to distinguish between 'true' zeros and 'protest' answers. Those who stated that they did not want to contribute because 'The other programme is more valuable', 'I can't afford it', 'I'm not concerned by this programme' were considered as true zeros. Those who gave other reasons were classified as protestors (see Protière et al. 2004 and Olsen et al. 2005). In cases of a protest answer for only one of the two programmes in the JE, the following rule has been applied. When there is a true zero and a protest answer, the respondent is considered as a protestor; when there is a positive value for one programme and a protest reason for the other, the missing value was replaced by a "zero", on the assumption that, if the respondent has given a value to one programme, s/he could not be seen as a protestor and then could not be excluded from the sample.[3].

All respondents who were willing to contribute were asked the following question: 'Knowing that your contribution would reduce what you have left to spend on other things, how much would you be willing to contribute each year for this programme?' Because it was a phone survey and to facilitate answers, the following pre-coded scale was used: less than 10€, 10 to 30€, 31 to 50€, 51 to 75€, 76 to 100€, 101 to 150€, and more than 150€. The middle of each range was used to represent a respondent's value.

In order to avoid any substitution or complementary effect in the JE exercise, before eliciting WTP for the second programme, respondents were informed that "the first programme was no longer a candidate for funding", thus rendering the programmes mutually exclusive.

Questions about respondents' socio-demographic characteristics, health behaviors,

[3] As in Luchini et al. (2003) a simultaneous estimation can be performed only on complete data set.

experience of cancer in general and of colorectal cancer in particular for members of family as well as own experience of screening were included in the questionnaire. Three scores are derived from some of these questions: two behavior scores and a score about the perception that respondents have of their own level of information toward factors that could influence health. "Healthy Attitude" and the "Non Healthy Attitude" scores (ScoreHB and ScoreNHB respectively) have been constructed on the base of 13 questions about (un)healthy lifestyle. That is to say, the scales were based on some objective factors that would reduce (raise) cancer risk, such as behaviors toward tobacco, alcohol (frequency, number of glasses, ...), sun (use of cream, cap, exposition hours), physical activity and nutrition. For the healthy attitude (resp. non healthy attitude) score, one point was allocated for each answer that reflected the more (resp. less) healthy attitude .

The "Non Informed" score (ScoreNI) was built with five questions about perception of information on alcohol effect, tobacco effect, cancer, alimentation and risk linked with sun exposition on health. For these five points the question was 'Do you feel you are very well; rather well; rather badly or very badly informed?'. To build the score, one point was allocated when the answer was 'badly' or 'very badly informed'.

4 Econometric model

Consider first the WTP of respondents in the SE exercise. In this case, the WTP W^s_{ij} of respondent i, $i = 1 \ldots n$, for programme j, $j = A, B, C$, is defined

$$W^s_{ij} = X_i\beta_j + \epsilon^s_{ij} \tag{5}$$

where X_i is a vector of explanatory variables, β_j a vector of parameters specific to programme j and ϵ_{ij} a well-behaved error term with mean zero and standard error σ^s_j that accounts for heterogeneity of tastes in the population.

In the joint evaluation case, the respondent evaluates sequentially, in the same questionnaire two health programmes j and k. S/he states a willingness to pay for each of the programmes W^d_{ij} and W^d_{ik}. As these willingness to pay values are elicited in the same CV exercise, we consider the following system of equations

$$\begin{cases} W^d_{ij} & = W^s_{ij} & + & \gamma^k_j & + & \epsilon^d_{ij} \\ W^d_{ik} & = W^s_{ik} & + & \gamma^j_k & + & \epsilon^d_{ik} \end{cases} \tag{6}$$

These equations state that the wilingness to pay in a joint evaluation context is a combination between that which would be obtained in a single valuation framework and specific effects induced by the joint valuation procedure. Typically, the description of another health programme in the evaluation exercise may influence a respondent's answer to the other programme and willingness to pay values can differ in the single and joint evaluation exercise. Namely, the parameter γ_j^k (resp. γ_k^j) stands for the impact on respondents' valuations of programme j (resp. k) when he also evaluates programme k (resp. j). Hereafter, these effects are called informational effects.

One might also consider that some sequence effects can occur. In our setting, the evaluation task required respondents to state their WTP for a pair of programmes j and k in different orders, either j first and k second (system 6) or k first and j second. In the latter case, the system of equations becomes

$$\begin{cases} W_{ik}^d = W_{ik}^s + \gamma_k^j + \delta_k^{kj} + \epsilon_{ik}^d \\ W_{ij}^d = W_{ij}^s + \gamma_j^k + \delta_j^{kj} + \epsilon_{ij}^d \end{cases} \tag{7}$$

where δ_k^{kj} and δ_j^{kj} stand for the impact on WTP of evaluating programme k first and programme j second (referent is evaluating programme j first and programme k second).

With three programmes A, B and C, the effects on WTP values of evaluating two programmes in JE instead of SE can be summarized as follows:

$$\begin{cases} \theta_A = \gamma_A^B + \gamma_A^C + \delta_A^{BA} + \delta_A^{CA} \\ \theta_B = \gamma_B^A + \gamma_B^C + \delta_B^{BA} + \delta_B^{CB} \\ \theta_C = \gamma_C^A + \gamma_C^B + \delta_C^{CA} + \delta_C^{CB} \end{cases} \tag{8}$$

The first two parameters in each equation stand for the impact on WTP values of evaluating jointly a second programme (with respect to evaluating only one programme in SE). The last two parameters stand for sequence effects that may occur in a JE that deals with pairs of programmes. It is then possible to define two nested hypothesis. First, when the informational effects are accounted for, thanks to the γ parameters, we can test for the presence of additional sequence effects by testing the nullity of the δ parameters:

$$\mathcal{H}_1 \text{ (no sequence effects):} \quad \delta_A^{BC} = \delta_A^{CA} = \delta_B^{BA} = \delta_B^{CB} = \delta_C^{CA} = \delta_C^{CA} = 0$$

A joint statistical test of nullity of this hypothesis and the γ parameters allows us to test our main assumption, that is SE and JE yield identical WTP values:

$$\mathcal{H}_0 \text{ (no difference between SE and JE): } \mathcal{H}_1 \text{ and } \gamma_A^B = \gamma_A^C = \gamma_B^A = \gamma_B^C = \gamma_C^A = \gamma_C^B = 0$$

It remains us to define the properties of the error terms ϵ_A^d, ϵ_B^d and ϵ_C^d. The standard error of each of the error terms is defined as follows:

$$\begin{cases} \sigma_A^d & = \sigma_A^s & + & \alpha_A \\ \sigma_B^d & = \sigma_B^s & + & \alpha_B \\ \sigma_C^d & = \sigma_C^s & + & \alpha_C \end{cases} \tag{9}$$

where σ_A^S, σ_B^S and σ_C^S are the standard errors computed when WTP values are obtained in SE and α_A, α_B and α_C are additional parameters introduced when programmes are evaluated in JE. Such a specification is two-fold. First it allows us to test for the different variances of WTP values in SE and JE. Second, it allows for some form of heteroscedasticity induced by estimated together WTP values elicited in SE and JE that, if not taken into account, may undermine statistical inference on our main hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$.

Finally, we consider the potential correlations that may occur between error terms in JE by specifying the following covariance matrix:

$$\Sigma_{jk} = \begin{bmatrix} (\sigma_j^d)^2 & \sigma_{jk} \\ \sigma_{jk} & (\sigma_k^d)^2 \end{bmatrix} \tag{10}$$

where σ_{jk}, $j = A, B, C$, $k = A, B, C$, is a covariance parameter that accounts for stochastic correlation between W_{ij}^d and W_{ik}^d. Under the assumption that the error terms are normally distributed, the model can be estimated by maximum likelihood using a bivariate normal distribution (three indeed, one for each pair of programmes). Note that the variance-covariance matrix of one particular contribution to the loglikelihood depends on which combination jk the respondent had to evaluate. For instance, the contribution to the likelihood of a respondent who performed a joint evaluation of programme A first and B second is:

$$L_i^{AB} = \phi\big(w_{iA} - X_i\beta_A - \gamma_A^B, w_{iB} - X_i\beta_B - \gamma_B^A, \Sigma_{AB}\big) \tag{11}$$

where ϕ is the density of the normal bivariate distribution and Σ_{AB} the variance covariance matrix of the joint evaluation (A, B). When the respondent performed a single evaluation exercise, his contribution to the likelihood is simply $L_i^j = \phi(w_{ij} - X_i\beta_j, \sigma_j^2)$ with $j = A, B, C$. Where ϕ in this case is the density of the univariate normal distribution.

Evaluation exercise		Programme A	Programme B	Programme C
A	Mean WTP (SD)	28.57 (±31.8)	-	-
$n = 260$	Nb of true zero	24 (9.2%)		
	Nb of prostestors	105 (40.4%)		
B	Mean WTP (SD)	-	28.45 (±35.5)	-
$n = 291$	Nb of true zero		43 (14.8%)	
	Nb of protestors		84 (28.9%)	
C	Mean WTP (SD)	-	-	32.46 (±34.8)
$n = 281$	Nb of true zero			29 (10.3%)
	Nb of protestors			96 (34.2%)
AB	Mean WTP (SD)	34.55 (±37.4)	30.16 (±33.9)	-
$n = 487$	Nb of true zero	56 (11.5%)	73 (14.9%)	
	Nb of protestors	121 (24.8%)	121 (24.8%)	
BA	Mean WTP (SD)	35.4 (±36.6)	30.62 (±37.9)	-
$n = 463$	Nb of true zero	40 (8.6%)	93 (20.1%)	
	Nb of protestors	118 (25.5%)	118 (25.5%)	
BC	Mean WTP (SD)	-	32.98 (±36.3)	32.44 (±35.3)
$n = 466$	Nb of true zero		56 (12.0%)	52 (11.2%)
	Nb of protestors		134 (28.8%)	134 (28.8%)
CB	Mean WTP (SD)	-	34.59 (±39.1)	35.36 (±38.7)
$n = 492$	Nb of true zero		58 (11.8%)	57 (11.6%)
	Nb of protestors		137 (27.8%)	137 (27.8%)
AC	Mean WTP (SD)	35.17 (±36.6)	-	26.87 (±32.3)
$n = 472$	Nb of true zero	32 (6.8%)		69 (14.6%)
	Nb of protestors	133 (28.2%)		133 (28.2%)
CA	Mean WTP (SD)	34.06 (±35.8)	-	28.94 (±34.9)
$n = 458$	Nb of true zero	43 (9.4%)		76 (15.7%)
	Nb of protestors	114 (24.9%)		114 (24.9%)

Table 1: WTP statistics, true zeros and protests by evaluation exercise ($n = 3670$)

5 Empirical results

Among the 4046 respondents, 226 (5.6%) had been cured for a cancer and were excluded from the survey. 150 (3.7%) additional respondents were excluded because they did not complete the questionnaire. The table in Appendix B presents characteristics of the 3670 respondents and a comparison between the 2628 (71.6%) contributors and the protestors. A number of significant differences could be observed between contributors and protestors. However no differences were found in terms of education, income, or experience of cancer. Some of the significant differences could be seen as not surprising, in particular three variables that could be seen as proxies of a high concern with health and of having some form of private protection, such as having subscribed to private health insurance, being vaccinated against grip and having a higher score on the healthy attitude scale. These respondents could be less willing to contribute to publicly funded programmes because they already perceive themselves as having greater levels of protection. A similar explanation could run for the higher proportion of respondents having experimented an hemocult test,[4] among protestors, in the sense that some of them could feel less concerned by programmes about colorectal cancer with which they have already experimented.

Table 1 provides descriptive statistics on WTP values, numbers of true zero and protestors according to the evaluation exercise: 832 (22.7%) respondents were included in the SE exercise and 2838 in the JE exercise. The main first observation could be that for programme A (screening) and B (treatment) the WTP values are higher in JE exercise than in the SE, whatever is the combination. On the other hand, WTP values for programme C (rehabilitation) are quite similar when evaluated alone or with treatment B but lower in SE then when evaluated with programme A.

For the sake of readability, econometric results are presented into two separate tables. Table 2 presents the parameter estimates associated with the covariates that explain WTP values for each of the three programmes. Table 3 presents the coefficients more specifically dedicated to the JE study. Consider first the parameter estimates associated with respondents' characteristics. Table 2 shows us that four variables are significant for all three programmes: *being a male* (MALE), *income* (INCOME), *having a member of the family with an experience of cancer* and *having already done an hemocult test* (EXPHEMO). All variables have a positive (but different) impact on WTP values. Some other variables have a significant positive impact on WTP values for programmes A and B only: *healthy behavior score* (ScoreHB), *Non informed Score* (ScoreNI) and *having a family member with cancer* (FAMILYK). Lastly, some variables are specific to one programme only. *Being a smoker* (SMOKER) have a positive impact on WTP values for programme A. Two variables are significant and have a positive impact for programme B only: *living in a rural area* (RURAL) and *having an University degree* (EDUC4).

Table 3 presents the parameter estimates associated with JE. The first three rows presents the covariance parameters between programmes. They are all highly significant and positive. This indicates that the independence assumption is not sustainable. Morever, the positiveness of these covariance coefficients means that the higher the value attached to one programme in a combination the higher will be the value for the other programme. The variance effects are also all highly significant and positive. Variances of WTP values are therefore significantly higher when elicited in JE rather than in SE. This effect is however less important for programme C.

[4]The hemocult test is a screening test for colorectal cancer.

Variable	Parameter estimates (p-value)		
	PROGRAMME A	PROGRAMME B	PROGRAMME C
Constant	11.0660	7.0887	21.8582
	(0.138)	(0.338)	(0.007)
MALE	3.2292**	5.4650***	4.5580**
	(0.032)	(0.000)	(0.012)
AGE	-0.2900	-0.3753	-0.4212
	(0.304)	(0.153)	(0.148)
AGE^2	0.0024	0.0035	0.0040
	(0.459)	(0.214)	(0.218)
INCOME	0.0047***	0.0037***	0.0033***
	(0.000)	(0.000)	(0.000)
EMPL	2.7412	0.5593	2.1572
	(0.192)	(0.793)	(0.306)
NBPERS	-0.6361	0.2604	0.3814
	(0.442)	(0.754)	(0.667)
NB16	1.1862	1.0732	0.5097
	(0.415)	(0.433)	(0.698)
HERED	5.5099**	6.6288***	3.8961**
	(0.014)	(0.001)	(0.150)
FAMILYK	2.5853*	2.8172*	2.2690
	(0.075)	(0.071)	(0.163)
SMOKER	4.5023***	1.9197	2.6879
	(0.009)	(0.298)	(0.137)
ScoreHB	1.0911***	1.0547**	0.4714
	(0.010)	(0.023)	(0.270)
ScoreNHB	-0.4059	1.0187*	-0.2962
	(0.532)	(0.098)	(0.653)
scoreNI	1.4839***	1.0687**	0.5323
	(0.007)	(0.042)	(0.376)
MUTUELLE	-3.1692	-0.1531	-0.1039
	(0.191)	(0.962)	(0.971)
EXPHEMO	7.1765***	7.6852***	4.7652*
	(0.002)	(0.000)	(0.055)
VGRIPPE	-1.0838	-3.2871	-1.2874
	(0.607)	(0.134)	(0.555)
RURAL	2.1936	4.6081***	1.4564
	(0.165)	(0.004)	(0.411)
EDUC2	-1.1811	-1.6083	-3.3227
	(0.576)	(0.462)	(0.145)
EDUC3	-0.2637	1.6115	-0.4686
	(0.902)	(0.481)	(0.840)
EDUC3	2.2011	5.9967***	0.2819
	(0.475)	(0.030)	(0.928)
Std Error σ_j	17.1200	20.7268	32.7514
	(0.000)	(0.000)	(0.000)
Loglikelihood		22231.608	

*** if p-value<0.01, ** if p-value<0.05, * if p-value<0.1

Table 2: Econometric results - covariates coefficients ($n = 2628$)

The last six lines of Table 3 presents the parameter estimates for the informational

Coefficients		Parameter estimates	p-value
Covariances			
σ_{AB}		1501.039***	(0.000)
σ_{AC}		1405.553***	(0.000)
σ_{BC}		1379.447***	(0.000)
Variance effects			
α_A		26.7940***	(0.000)
α_B		21.9441***	(0.000)
α_C		6.4142***	(0.001)
Informational effects			
on WTP for programme A when elicited with	programme B: γ_A^B	6.2016**	(0.022)
	programme C: γ_A^C	6.2735**	(0.024)
on WTP for programme B when elicited with	programme A: γ_B^A	2.1210	(0.426)
	programme C: γ_B^C	4.8213*	(0.081)
on WTP for programme C when elicited with	programme A: γ_C^A	-4.2954	(0.182)
	programme B: γ_C^B	1.5059	(0.668)
Sequence effects			
on WTP for programme A when elicited in	sequence BA: δ_A^{BA}	0.7807	(0.851)
	sequence CA: δ_A^{CA}	0.6065	(0.879)
on WTP for programme B when elicited in	sequence BA: δ_B^{BA}	0.2390	(0.957)
	sequence CB: δ_B^{CB}	2.8903	(0.434)
on WTP for programme C when elicited in	sequence CA: δ_C^{CA}	1.7621	(0.656)
	sequence CB: δ_C^{CB}	2.4509	(0.471)

*** if p-value<0.01, ** if p-value<0.05, * if p-value<0.1

Table 3: Econometric results - JE effects ($n = 2628$)

effects and the sequence effects potentially induced by JE. Consider first the informational effects. Three of them are significant and positive: γ_A^B ($p = .022$), γ_A^C ($p = .024$) and γ_B^C, although less significantly ($p = .081$). WTP for the screening programme A is therefore more likely to be higher when elicited with programme B or C. Similarly, WTP for the treatment programme B is more likely to increase when jointly evaluated with the rehabilitation programme C. On the contrary, eliciting jointly programme B together with the screening programme A or C with another programme, whatever it is, does not yield different WTP values in comparison with SE. Concerning the sequence effects, that is δ coefficients, none of them pass the significance test individually.

We finally focus on the tests of our main assumptions $\mathcal{H}_0$ and $\mathcal{H}_1$. We do so by estimating two restricted models in order to test for the presence of sequence effects $\mathcal{H}_1$ and the absence of any mean effect induced by the joint evaluation procedure $\mathcal{H}_0$.[5]

[5]These constrained models keep with different variances between SE and JE. Given the high signif-

This allows us to compute LR tests. Results are two-fold. First, we cannot reject the null hypothesis $\mathcal{H}_1$ of no sequence effect: LR=3.20 with p-value=.800. Second, the LR test for the complete model against the null hypothesis $\mathcal{H}_0$ of no differences between SE and JE is: LR=16.45 with p-value<0.0001. We therefore reject the null of no difference between SE and JE. In our data, we thus conclude that, although the JE exercise leads to different WTP values due to informational effects, it does not induce sequence effects.

6 Discussion

Two main aspects of our results are expected and, to an extent, confirm the validity of the exercise conducted. These are the relationships between income and WTP and also those with cancer experience who were willing to pay registering higher WTP values.

With respect to the main issue of SE versus JE, we can say three things. First, there is a difference between WTP values that arise in the SE and JE situations. This is, in itself, an important matter of concern for public decision making. Second, the econometric results have shown that this is due neither to heterogeneity in respondents nor to sequencing effects. Third, and following on from the second point, these results are important because they highlight the main possibility of differences between JE and SE being due to informational effects; in line with economic theory, it could be argued that this is because respondents were provided with infomation on another alternative when making their evaluation in the JE situation. Looking back at programme A relative to B as an example, in the SE situation it might be that respondents simply register a 'vote' for each of A and B relative to a reference point which is either random or zero, resulting in similar WTP values. In the JE situation, respondents see the relevant competing programmes and correspondingly provide more discriminating WTP values. The isolation of informational effects has not been shown before, previous studies mixing such effects with those arising from sequencing.

Of course, the example provided in the previous paragraph is speculative. Although we, as well as others, have argued in favour of a JE approach as a way of 'completing the market', more systematic research is required in order to assess the reproducability and relevance of the informational effects we have isolated.

icance of the additional variance terms α_A, α_B and α_C in JE, we would obviously reject a constrained model with identical variances in SE and JE and therefore not be able to disantangle variance and mean effects induced by JE.

7 Conclusion

The shortcomings of CV when measured against the market are well acknowledged and have received much attention, focussing on issues such as hypothetical bias for example. Innovatory methods, such as 'cheap talk' (Cummings and Taylor 1999; Aadland and Caplan 2006; K. et al. 2008) and group-evaluation procedures (Alvarez-Farizo and Hanley 2006) have been devised, and with some success, to alleviate such challenges. These innovations are not mutually exclusive, however. In this study, we have demonstrated the potential for JE approaches to be added to this armoury in a way that induces informational effects without incurring problems of sequencing.

References

Aadland, David and Arthur J. Caplan (2006). Cheap talk reconsidered: New evidence from CVM. *Journal of Economic Behavior & Organization*, **60**, 562–578.

Alvarez-Farizo, B. and N. Hanley (2006). Improving the process of valuing non-market benefits: combining citizens' juries with choice modelling. *Land Economics*, **82**, 465–478.

Bazerman, M., D. Moore, A. Tenbrunsel, K. Wade-benzoni, and S. Blount (1999). Explaining how preferences change across joint versus separate evaluation. *Journal of Economic Behaviour and Organization*, **39**, 41–58.

Carson, R., N.E. Flores, and W.M. Haneman (1998). Sequencing and valuing public goods. *Journal of Environmental Economics and Management*, **36**, 314–323.

Cummings, Ronald G. and Laura O. Taylor (1999). Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review*, **89**, 649–665.

Donaldson, C., P. Schackley, M. Abdalla, and Miedzybrodzka (1995). Willingness to pay for antenatal screening for cystic fibrosis. *Health Economics*, **4**, 439–452.

Donaldson, C., P. Shackley, and M. Abdalla (1997). Using willingness to pay to value close substitutes: Carrier screening for cystic fibrosis revisited. *Health Economics*, **6**, 145–159.

Ferlay, J., E. Bray, P. Pisani, and D.M. Parkin (2004). Globocan 2002: cancer incidence, mortality and prevalence worldwide. IARC CancerBase number 5, version 2.0, IARC Press, Lyon, France.

Guilbert, P., P. Peretti-Watel, F. Beck, and A. Gautier (2006). Baromètre cancer 2005. Saint-Denis: INPES, Technical report.

Hoehn, J.P. and A. Randall (1989). Too many proposals pass the benefit cost test. *American Economic Review*, **70**, 544–551.

Hsee, C.K. (1996). The evaluability hypothesis: an explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behaviour and Human Decision Processes*, **67**, 247–245.

Hsee, C.K. (1998). Less is better: when low-value options are valued more highly than high value options. *Journal of Behavioural Decision Making*, **11**, 107–121.

Johannesson, M. and B. Fegeberg (1992). A health-economic comparison of diet and drug treatment in obese men with mild hypertension. *Journal of Hypertension*, **10**, 1063–1070.

K., Blumenschein, G.C. Blomquist, M. Johannesson, N. Horn, and P. Freeman (2008). Eliciting willingness to pay without bias: evidence from a field experiment. *Economic Journal*, **118**, 114–137.

Kahneman, D. and I. Ritov (1994). Determinants of stated preference for public goods: a study in the headline method. *Journal of Risk and Uncertainty*, **9**, 5–38.

Luchini, S., C. Protière, and JP. Moatti (2003). Eliciting several willingness to pay in a single contingent valuation survey: application to health care. *Health Economics*, **12**, 51–64.

Olsen, JA. and C. Donaldson (1998). Helicopters, hearts and hips: Using willingness to pay to set priorities for public sector health care programmes. *Social Science and Medicine*, **46**, 1–12.

Olsen, JA., C. Donaldson, P. Shackley, and the EuroWill Group (2005). Implicit versus explicit ranking: on inferring ordinal preferences for health care programmes based on differences in willingness-to-pay. *Journal of Health Economics*, **24**, 990–996.

Payne, J. W., D. A. Schkade, W.H. Desvouges, and C. Aultman (2000). Valuation of multiple environmental programs. *Journal of Risk and Uncertainty*, **21**, 95–115.

Protière, C., C. Donaldson, J.P. Moatti, S. Luchini, and P. Shackley (2004). The impact of information on non-health attributes on willingness to pay for multiple health care programmes. *Social Science & Medecine*, **58**, 1257–1269.

Ryan, M., J. Ratchliffe, and J. Tucker (1997). Using willingness to pay to value alternative models of antenatal care. *Social Science and Medicine*, **44**, 371–380.

Stewart, J., E. O'Shea, C. Donaldson, and P. Shackley (2002). Do ordering effects matter in willingness to pay studies of health care? *Journal of Health Economics*, **21**, 585–599.

A Scenarios

"We are going to talk about colorectal cancer. It's the most frequent cancer. It occurred at 70 years old in average and is at the origin of more than 16 000 death by year. This cancer could be hereditary, but is also favored by a too rich alimentation or a lack of exercise."

The "screening" programme A

A screening made every two years among the more than 50 years old would allow to decrease the mortality rate by colorectal cancer of 18% and to avoid some radiotherapy and chemotherapy to patients. This screening consists in the search of blood in selle, and if necessary followed by a coloscopy

The "new treatment" programme B

A new kind of chemotherapy allows to reduce the cases of death or relapse from 26 to 21 persons on 100. However this new chemotherapy is more toxic and could induce an hospitalization in nearly 10 on 100 patients.

The "rehabilitation" programme C

The aim of this programme is to help rehabilitation of people who have been cured for a colorectal cancer by making available a ?home aid? and by reimbursing prosthesis and accessories needed to have a "normal life" (wig, ...)

B Sample characteristics

Variable name	Variables description (n=2628)	Contributors (n=1042)	Protestors (n=3670)	Total
SEX	Male **	1102 (41.9%)	398 (38.2%)	1500 (40.9%)
AGE	Mean age (SD) **	42.35 (17.9)	49.05 (16.6)	44.25 (17.8)
NBPERS	Number of person in household (SD) **	2.73 (1.4)	2.48 (1.3)	2.66 (1.4)
NB16	Number of person in household aged 16 or over (SD) **	2.10 (0.9)	1.94 (0.8)	2.05 (0.9)
DIP1	No diploma. Primary school. Professional certificate (SD)	1370 (52.1%)	546 (52.4%)	1916 (52.2%)
DIP2	Secondary	451 (17.2%)	173 (16.6%)	624 (17.0%)
DIP3	High school certificate	565 (21.5%)	229 (22.0%)	794 (21.6%)
DIP4	University degree	236 (9.0%)	90 (8.6%)	326 (8.9%)
MDEALT	Occupational status in medical environment	175 (6.7%)	58 (5.6%)	233 (6.3%)
EMPL	Being in professional activity **	1389 (52.9%)	604 (58.0%)	1993 (54.3%)
INCOME	Mean household monthly income €(SD)	2083 (1240)	2107 (1202)	2090 (1229)
RURAL	Living in a rural area	758 (28.9%)	313 (30.2%)	1071 (29.2%)
MUTUELD	Has subscribed a private (for-profit or not-for-profit) health insurance in addition to Social Security coverage**	2409 (91.7%)	983 (94.3%)	3392 (92.4%)
VACGRIP	Being vaccinated against grippe**	532 (20.2%)	256 (24.6%)	788 (21.5%)
SMOKER	Being a smocker	1246 (47.4%)	463 (44.4%)	1709 (46.6%)
ScoreHB	Healthy Behavior Score (0 to 13)*	7.32 (2.4)	7.47 (2.4)	7.36 (2.4)
ScoreNHB	Non Healthy Behavior Score (0 to 13)**	1.91 (1.5)	2.05 (1.5)	1.95 (1.5)
ScoreNI	Non informed Score (0 to 5)*	1.01 (1.3)	1.11 (1.4)	1.04 (1.3)
FAMILYK	Having member of family with cancer	1501 (57.1%)	589 (56.5%)	2090 (56.9%)
HERED	Having a parent or a grand parent who had a colorectal cancer	273 (10.4%)	106 (10.2%)	379 (10.3%)
EXPHEMO	Having experimented an Hemoccult test **	342 (13%)	174 (16.7%)	516 (14.1%)